Posudek oponenta diplomové práce

Jméno a příjmení autora posudku: *Petr Gregor*

Jméno a příjmení autora práce: *Filip Bártek*

Název práce: *Minimální reprezentace víceintervalových booleovských funkcí*

Booleovská funkce $f\colon \{0,1\}^n \rightarrow \{0,1\}$ *je* k-intervalová, je-li množina jejích true-pointů sjednocením k intervalů lexikografického uspořádání množiny $\{0,1\}^n$. Diplomová práce se zabývá hledáním minimální DNF reprezentace k-intervalových funkcí, tedy (přesného) pokrytí jejich true-pointů co nejmenším počtem (ne nutně disjunktních) podkrychlí n-dimenzionální hyperkrychle. Cílem práce bylo pokusit se rozšířit výsledky Schiebera et al. z článku [1] o 1-intervalových funkcích na víceintervalové funkce a navázat na předchozí diplomovou práci Dubovského [2] o 2-intervalových funkcích. Bylo dosaženo těchto výsledků:

1. zjednodušený algoritmus minimalizace pro 2-zlomové 2-intervalové funkce oproti algoritmu v [2],
2. existence tzv. *non-coverable* l-zlomové funkce arity $l+1$ pro každé $l>2$, což v důsledku znamená, že dosavadní technika tzv. *ortogonálních* množin se nedá použít k důkazu optimality algoritmů pro více než 2-zlomové 2-intervalové funkce,
3. $2k$-aproximační algoritmus pro DNF reprezentaci k-intervalových funkcí založený na sufix-prefix dekompozici,
4. komplikovanější algoritmus pro tentýž problém s aproximačním poměrem mezi $2k-2$ a $2k$.

Uvedené výsledky jsou dle mého soudu správně, až na použití špatné funkce $f^4_{[0,4],[11,14]}$ v důkazu Theorem 3.2.1. Ta má totiž ortogonální množinu $\{2,4,11,13,14\}$ velikosti 5 a zároveň pokrytí pomocí 5 podkrychlí. Místo ní je třeba vzít funkci $f^4_{[0,4],[9,14]}$, o níž Dubovský [2] ukázal, že není coverable. Jde tedy jen o snadno opravitelné přehlédnutí, i když třikrát se opakující (str. 23,25).

Dle mého názoru není práce příliš originální, nepřináší nové přístupy, nové téma či otázky, jen zjednodušuje či přímočaře zobecňuje dosavadní výsledky. Navržené aproximační algoritmy jsou (triviálním) použitím známých algoritmů pro 1-intervalové funkce separátně na jednotlivé intervaly. Na druhou stranu je třeba říct, že výsledky jsou v souladu se zadáním práce. Oceňuji také, že autor pečlivě cituje předchozí práce a je tedy vždy jasné, jaký je jeho vlastní přínos.

Práce je psaná přehledně s angličtinou na velmi dobré úrovni, jen s malým množstvím chyb. Občas se nevyhne některým neobratnostem, např. zavádí zvlášť termíny, že ternární vektor „*spans / spans exactly*“ binární vektor / množinu vektorů, ale v Lemma 2.2.3 a jeho důkazu se mezi těmito termíny nerozlišuje, a používá se „*spans*“ ve smyslu „*spans exactly*“ (str. 14). Nebo Observation 2.4.1. jistě není jen pozorování, jde o zásadní důsledek hlubšího výsledku [1, Theorem 3].

Za jeden z hlavních nedostatků práce ale považuji to, že je v několika ohledech neúplná:

- Na několika místech se okrajově zmiňuje složitost studovaných algoritmů v pojmech jako „*efficient*“, „*simple*“, „*performs better*“, „*runs in polynomial time*“. Nikde se ale časová složitost vzhledem k parametrům n a k nezmiňuje. Přeci jen bych u teoretické práce, jež nejméně z poloviny navrhuje a studuje nějaké algoritmy, očekával (nějaký) rozbor jejich složitost.
- Ačkoliv jsou všechny algoritmy popsané úplně (self-contained), důkaz optimality v sekci 3.1.3 popsaný úplně není, opírá se totiž o ortogonální množiny z článku [1], které v diplomové práci popsané nejsou.
- Chybí analogické pozorování jako Observation 5.2.11 pro true-pointy, které končí $1^{(n-k)}$, ačkoliv se (implicitně) používá (na str. 37, 38 dole). Jde ale spíš jen o opomenutí.
- Postrádal jsem i bližší vysvětlení proč studovat intervalové funkce vzhledem k právě lexikografickému uspořádání, když je možné si představit i jiná uspořádání. Například se zcela pomíjí možnost zkoušet permutace souřadnic. Je snadné najít k-intervalovou funkci pro libovolné k, která se po vhodné permutaci souřadnic stane 1-intervalovou. Taková je i funkce použitá v důkazu Theorem 4.3.3 pro dolní odhad aproximačního poměru. Permutace souřadnic přitom zachovávají ortogonální množiny.

Za druhý hlavní nedostatek považuji to, že Algoritmus 4.1 (sufix-prefix dekompozice) není přímočarým rozšířením algoritmu z článku Schieber et al. [1, Section 6], jak se v práci zmiňuje (str. 26 nahoře), ale jeho mnohem horší varianty. Přesněji, očekával bych, že pro $k=1$ dosáhne Algoritmus 4.1 stejných výsledků jako algoritmus z [1, Section 6], je ale horší. Stačí porovnat, jak se algoritmy zachovají na vstupu tvaru $a=c0^{(n-j)}$, $b=c1^{(n-j)}$. V prvním případě dojde k dekompozici a dostaneme dva ternární vektory, v druhém pouze jeden. Svědčí o tom i porovnání aproximačních poměrů v disjunktním případě, viz Theorem 4.3.3. a [1, Theorem 9]. Nerozumím, proč byl Algoritmus 4.1 navržen tak nevhodně, jak je.

I přes uvedené nedostatky je práce celkově na dobré úrovni a doporučuji ji uznat jako diplomovou práci.

[1] B. Schieber, D. Geist, A. Zaks, Computing the minimum DNF representation of Boolean functions defined by intervals, *Discrete Applied Mathematics* **149** (1-3), 154-173, 2005.

[2] J. Dubovský, A construction of minimum DNF representations of 2-interval functions, Master's thesis, Charles University in Prague, 2012.

Doporučení k obhajobě:

Z výše uvedených důvodů práci *doporučuji* k obhajobě.

Vynikající práce vhodná pro soutěž studentských prací:	NE

V Praze dne: 27. 5. 2015

Podpis: Petr Gregor, v.r.